A ROLE FOR HOMOSEXUALITY IN POPULATION GENETICS

Marion Cook
En Vignette, Echenevex, 01170 Gex, France

Received 7 October 1993, 14 December 1995

ABSTRACT. In this paper it is shown that a hypothetical gene for male homosexuality would be rapidly lost from a population, and that recurrent mutation would be necessary for its maintenance. A short-term advantage in terms of population dynamics, dependent on the reduction in variance through time in mean fitness, arises from this very fact. The population is more stable in size, and can have a higher long-term growth rate in a fluctuating environment, compared to an all heterosexual population. The concern that precise identification of a gene for homosexuality could result in interference with the genetic code, or the abortion of potentially homosexual foetuses, may thus be unfounded. Such measures, apart from being unethical, would in all probability be ineffectual in lowering the frequency of homosexuality if recurrent mutation were involved.

* * *

Recent studies have indicated a genetic component in human homosexuality[1,2]. In particular Hamer et al[3,4] have identified a correlation between male homosexual orientation and the inheritance of polymorphic markers on the *X* chromosome in the *Xq*28 region. This suggests that one form of male homosexuality might be passed on through the female line.

From the point of view of simple Darwinian theory, the presence of male homosexuality in a population, if genetically determined, constitutes an enigma[5] since a gene for homosexuality would be expected to be rapidly lost, homosexual males breeding less frequently than heterosexual males.

If a gene for homosexuality did exist, it could be maintained in the population by an overdominance effect or by recurrent mutation.

To maintain a gene for male homosexuality carried on the *X* chromosome through an overdominance effect[6] would entail a large fitness advantage to female carriers. The greatest advantage necessary would be if no homosexual males bred. In terms of fecundity, assuming a carrier has two young, then she has a half chance of having a son, and a half chance he will be homosexual, thus her fitness would be three-quarters that of a non-carrier, and she would need to have more progeny to be competitive.

No sign of greater fecundity of the mothers of homosexuals is recorded. In terms of survival, such an advantage to carriers is

* * *

Evolutionary Theory 11: 135-151 (May, 1996)
The editors thank S. Orzack and another referee for help in evaluating this paper.

unlikely to be hidden in the low mortality rates from infancy to middle age for the developed world. Even if such an advantage had existed in the past, any gene for homosexuality would have been lost, or would be extremely rare, by now.

That recurrent mutation is likely to be involved in the maintenance of a gene for male homosexuality in a population, whether carried on the *X* chromosome or an autosomal chromosome, is shown in the models to follow.

A short-term advantage in terms of population dynamics is demonstrated to result from the very fact that homosexual males breed less frequently[7]. The advantage depends on the reduction in variance through time in mean fitness of the population, and can be especially strong in a population where females are *XX* and males *XY* when a hypothetical gene for male homosexuality is borne on the *X* chromosome. The population is more stable in size and can have a higher long-term growth rate in a fluctuating environment compared to an all heterosexual population. Although recurrent mutation is necessary to maintain the hypothetical gene for homosexuality in the population, the advantage persists no matter how low the rate of mutation or the frequency of the gene.

It is thought by some (see ref. 8) that there may be no such thing as random or undirected mutation, mutations only being generated if and when needed.

Assuming recurrent mutation is necessary to maintain a gene for homosexuality in a population, then mutant carriers with the most surviving, i.e. fittest, offspring would have the best probability of passing on the gene.

Consider a newly mutated recessive gene, *h*, appearing in a diploid carrier, *Hh*, that must mate in a wild-type population. This population if fairly constant in size has a longterm geometric mean fitness, *LGMF*, of one, each mating between two individuals producing on average two surviving offspring to compensate for the eventual loss of the parents. Of course some matings are less and some are more fruitful, and there is a range in family size from zero to many. A family with many offspring has a higher probability of preserving the newly mutated gene than a family with few offspring[9]. The chance, *c*, of losing the mutant gene *h* in the filial generation in a mating between the mutant carrier, *Hh*, and a wild-type, *HH*, depends on fitness, *f*, in terms of surviving offspring:

$$c = \frac{1}{2^f} \qquad (1)$$

If no offspring are produced or survive the mutant gene *h* will be lost, if there is one surviving offspring the probability of *h* being lost is 1/2, if two the probability is 1/4, and so on.

The same probability of losing the mutant gene applies when it is the dominant allele, *H*, and the wild-type is *hh*.

Fisher[10] has shown that after *n* generations the probability of a single neutral gene surviving, in the Wright-Fisher model, is only about $2/n$ where *n* is large. The probability of a gene for homosexuality surviving would be even lower, since homosexuals are less likely to breed. The turnover rate of the homosexuality gene at the *H* locus in the population under recurrent mutation would be relatively rapid and the gene rare.

Consider now a large diploid population diallelic at each of two loci, the *H* locus with the rare homosexuality gene requiring recurrent mutation for its maintenance in the population, and the *A* locus, any locus coding for general viability, not requiring recurrent mutation under fluctuating selection in a two-state environment for the maintenance of polymorphism, since the polymorphism is protected. Individuals from families that are the fittest in terms of surviving offspring will tend to possess the currently fitter allele at the *A* locus. Since the turnover rate of the rare homosexuality gene at the *H* locus would be fast, carriers being mostly from large families would likewise tend to possess the currently fitter allele at the *A* locus.

At the *A* locus, each allele will tend to increase in frequency in the environment favouring it. This can cause fluctuations in population size. It has been found[11,12] that reduction in variance through time in mean fitness can lead to a higher *LGMF* of a population. If a proportion of males bearing the currently fitter allele at the *A* locus did not breed, the polymorphism could be buffered. A role for the rare homosexuality gene at the *H* locus might lie here, since homosexuals are less likely to breed and would tend to possess the currently fitter allele at the *A* locus.

A diallelic locus such as the *A* locus affecting viability has been modelled[13,14] (and see Models below). It was shown that where the fitness of a male in terms of offspring depended on the reproductive success of his mate there was additive variance in fitness in a two-state environment under fluctuating selection, the resulting polymorphism being protected. Indications were that the population fluctuated less in size when males from small families were more successful in attaining mates, and that this lower variance through time in mean fitness usually raised the *LGMF*[13]. This buffering effect could be provided by the *H* locus for the rare homosexuality gene if homosexuals tended to be from large families, since few breed.

The effect of a hypothetical gene for homosexuality on population fitness when carried on the *X* chromosome and on an autosomal chromosome was modelled using computer simulations written in Turbo Pascal.

The Models

In the models males are assumed to contribute only genes to the next generation. Mating is random, and there is no kin selection, sib competition or male competition. Although males equal females in frequency, homosexual males are less likely to reproduce. It is assumed that this has no effect on female reproductive success, some males having more than one mate and all females mating.

The population is assumed to be large. Generations are non-overlapping and discrete.

The two diallelic loci show complete dominance, there being no overdominance or partial dominance. The *A* locus determining viability has a protected genetic polymorphism under fluctuating selection in a two-state environment, while the *H* locus requires recurrent mutation to maintain the rare gene for homosexuality.

At the viability locus the dominant genotype *A-* confers a fitness (W_{bd} in the equations below) up to the time of mating on both males and females of b in environment one and b^{-1} in environment two, while the recessive genotype *aa* confers a fitness (W_{br}) of b^{-1} in environment one and b in environment two.

After mating the viability locus continues to affect reproductive success through its action on the female[14], and hence on her dependent young, *A-* conferring a fitness (W_{fd} in the equations below) of f in environment one and f^{-1} in environment two, *aa* a fitness (W_{fr}) of f^{-1} in environment one and f in environment two. Action of the viability locus on the male has no effect on his reproductive success after the time of mating since he contributes only genes in the models.

To simplify the argument and highlight the maternal effect it is assumed that genes at the viability, *A*, locus have no effect on the young before sexual maturity, i.e. $b = 1$. For each population trajectory, the value of f giving a $LGMF = 1$ is first computed using the simulation without mutation, u, at the *H* locus, i.e. in the absence of the gene for homosexuality by using the value $u = 0$ in the equations below. Initial allele frequencies at the *A* locus are not important since the polymorphism is protected[14].

At the homosexuality locus, mutation of the wild type gene to the homosexuality allele is assumed to be recurrent. The reverse mutation is ignored since it usually occurs at a lower rate and because the frequency of the homosexuality allele is so low. A constant mutation rate, u, is used throughout each computer run of a population trajectory.

The chance of the mutant gene not being passed on by a male or female carrier, *Hh*, to their offspring depends on maternal fitness in terms of surviving progeny, W_{fd} for an *A-* female, W_{fr} for an *aa* female, and is determined as in the equations:

$$c_d = 1 / 2^{N_{fd}} \quad (2a)$$

$$c_r = 1 / 2^{N_{fr}} \quad (2b)$$

c_d being the chance the mutant gene at the *H* locus will be lost if an *A*- female is the mother, and c_r the chance if an *aa* female is the mother. Where maternal fitness is f^{-1} it is assumed that only one young survives in the occasional successful case, the probability of losing the mutant gene thus being 0.5.

Population trajectories are computed over cycles of *n* generations until allele frequencies settle into a repeating sequence, usually in fewer than 200 generations. The conditions cycle through $0.5n$ generations in environment one, in which a female genotype sees a fitness of *f* at the viability, *A*, locus if the environment is favourable, or f^{-1} if unfavourable, and $0.5n$ generations in environment two, in which the female genotype experiences the alternate environment. The number of generations in each environment is considered identical to facilitate comparisons, but this is not a requisite. The proportion, *P*, of homosexual males breeding is the same in both environments.

Homosexuality Gene Locus on an Autosomal Chromosome

The effect of the *H* locus on population stability when carried on an autosomal chromosome, with the hypothetical gene for homosexuality recessive or dominant, is first investigated.

Homosexuality gene recessive

Homosexual males are of genotype *hh* in this model, *u* being the rate of mutation of *H* to *h*, and *P* is the proportion breeding.

The frequency of the gamete type *AH* = p_0, *Ah* = p_1, *aH* = p_2 and *ah* = p_3. These lead to the 10 genotype frequencies *AHAH* = G_{00}, *AHAh* = G_{01}, *AHaH* = G_{02}, *AHah* = G_{03}, *AhAh* = G_{11}, *AhaH* = G_{12}, *Ahah* = G_{13}, *aHaH* = G_{22}, *aHah* = G_{23}, and *ahah* = G_{33}.

Genotypes surviving to breeding time are:

$$\begin{aligned}
G'_{00} &= W_{bd}\, G_{00} / \overline{W_b} & G'_{12} &= W_{bd}\, G_{12} / \overline{W_b} \\
G'_{01} &= W_{bd}\, G_{01} / \overline{W_b} & G'_{13} &= W_{bd}\, G_{13} / \overline{W_b} \\
G'_{02} &= W_{bd}\, G_{02} / \overline{W_b} & G'_{22} &= W_{br}\, G_{22} / \overline{W_b} \\
G'_{03} &= W_{bd}\, G_{03} / \overline{W_b} & G'_{23} &= W_{br}\, G_{23} / \overline{W_b} \\
G'_{11} &= W_{bd}\, G_{11} / \overline{W_b} & G'_{33} &= W_{br}\, G_{33} / \overline{W_b}
\end{aligned} \quad (3)$$

where

$$\overline{W_b} = \Sigma\, W_b\, G_{ij} \quad (4)$$

and W_b is either W_{bd} or W_{br}, *b* = 1 being used for the figures.

The mutation episode for the female gives:

$$
\begin{aligned}
G_{11f} &= G'_{11} + uG'_{01} + u^2G'_{00} \\
G_{01f} &= (1-u)G'_{01} + (2u-u^2)G'_{00} \\
G_{00f} &= (1-2u)G'_{00} \\
G_{13f} &= G'_{13} + uG'_{03} + uG'_{12} + u^2G'_{02} \\
G_{03f} &= (1-u)G'_{03} + \tfrac{1}{2}(2u-u^2)G'_{02} \\
G_{12f} &= (1-u)G'_{12} + \tfrac{1}{2}(2u-u^2)G'_{02} \\
G_{02f} &= (1-2u)G'_{02} \\
G_{33f} &= G'_{33} + uG'_{23} + u^2G'_{22} \\
G_{23f} &= (1-u)G'_{23} + (2u-u^2)G'_{22} \\
G_{22f} &= (1-2u)G'_{22}
\end{aligned} \qquad (5)
$$

the subscript f labels female genotypes.

Gamete proportions for the female are:

$$
\begin{aligned}
p_{0fd} &= [2G_{00f} + (1+c_d)((1-r)G_{03f}+rG_{12f}+G_{01f}) + G_{02f}] / 2 \\
p_{1fd} &= [2G_{11f} + (1-c_d)((1-r)G_{12f}+rG_{03f}+G_{01f}) + G_{13f}] / 2 \\
p_{2fd} &= [G_{02f} + (1+c_d)((1-r)G_{12f}+rG_{03f})] / 2 \\
p_{3fd} &= [G_{13f} + (1-c_d)((1-r)G_{03f}+rG_{12f})] / 2 \\
p_{2fr} &= [2G_{22f} + (1+c_r)G_{23f}] / 2 \\
p_{3fr} &= [2G_{33f} + (1-c_r)G_{23f}] / 2
\end{aligned} \qquad (6)
$$

the subscripts fd and fr label gametes from A- and aa females respectively. Recombination, r, (0 indicating no crossing-over to 0.5) is included. The chance of the mutant gene h surviving is 1 - c; c values being obtained from equation (2).

Male genotypes breeding are:

$$S'_m = G'_{00} + G'_{01} + G'_{02} + G'_{03} + PG'_{11} + G'_{12} + PG'_{13} + G'_{22} + G'_{23} + PG'_{33} \qquad (7)$$

the subscript m labels male genotypes.

The mutation episode for the male gives:

$$
\begin{aligned}
G_{11m} &= [PG'_{11} + uG'_{01} + u^2G'_{00}] / S'_m \\
G_{01m} &= [(1-u)G'_{01} + (2u-u^2)G'_{00}] / S'_m \\
G_{00m} &= [(1-2u)G'_{00}] / S'_m \\
G_{13m} &= [PG'_{13} + uG'_{03} + uG'_{12} + u^2G'_{02}] / S'_m \\
G_{03m} &= [(1-u)G'_{03} + \tfrac{1}{2}(2u-u^2)G'_{02}] / S'_m \\
G_{12m} &= [(1-u)G'_{12} + \tfrac{1}{2}(2u-u^2)G'_{02}] / S'_m \\
G_{02m} &= [(1-2u)G'_{02}] / S'_m \\
G_{33m} &= [PG'_{33} + uG'_{23} + u^2G'_{22}] / S'_m \\
G_{23m} &= [(1-u)G'_{23} + (2u-u^2)G'_{22}] / S'_m \\
G_{22m} &= [(1-2u)G'_{22}] / S'_m
\end{aligned} \qquad (8)
$$

Gamete proportions for the male are:

$$
\begin{aligned}
p_{0md} &= [2G_{00m} + (1+c_d)((1-r)G_{03m}+rG_{12m}+G_{01m}) + G_{02m}] / 2 \\
p_{1md} &= [2G_{11m} + (1-c_d)((1-r)G_{12m}+rG_{03m}+G_{01m}) + G_{13m}] / 2 \\
p_{2md} &= [2G_{22m} + (1+c_d)((1-r)G_{12m}+rG_{03m}+G_{23m}) + G_{02m}] / 2 \\
p_{3md} &= [2G_{33m} + (1-c_d)((1-r)G_{03m}+rG_{12m}+G_{23m}) + G_{13m}] / 2
\end{aligned} \qquad (9)
$$

The chance of a male mutant carrier Hh passing on the gene h depends on the female he mates with, e.g. p_{1md} contains the chance of gamete Ah from the male being passed on if mating is with an A-female, p_{1mr} if mating is with an aa female; p_{0mr}, p_{1mr}, etc. being calculated as in (9) by replacing c_d with c_r. In the union of gametes equations (10) below the gamete frequency p_{1md} is used if fusion is with a p_{fd} type gamete and p_{1mr} if fusion is with a p_{fr} type gamete.

Next comes the random union of gametes. W_{fd} and W_{fr} are the fitness after mating of A- and aa females respectively, f values used depending on the environment. The surviving zygotes depend on maternal fitness after mating:

$$
\begin{aligned}
G''_{00} &= [W_{fd}\,p_{0fd}\,p_{0md}] / \overline{W_f} \\
G''_{01} &= [W_{fd}\,p_{0fd}\,p_{1md} + W_{fd}\,p_{1fd}\,p_{0md}] / \overline{W_f} \\
G''_{02} &= [W_{fd}\,p_{0fd}\,p_{2md} + W_{fd}\,p_{2fd}\,p_{0md} + W_{fr}\,p_{2fr}\,p_{0mr}] / \overline{W_f} \\
G''_{03} &= [W_{fd}\,p_{0fd}\,p_{3md} + W_{fd}\,p_{3fd}\,p_{0md} + W_{fr}\,p_{3fr}\,p_{0mr}] / \overline{W_f} \\
G''_{11} &= [W_{fd}\,p_{1fd}\,p_{1md}] / \overline{W_f} \\
G''_{12} &= [W_{fd}\,p_{1fd}\,p_{2md} + W_{fd}\,p_{2fd}\,p_{1md} + W_{fr}\,p_{2fr}\,p_{1mr}] / \overline{W_f} \\
G''_{13} &= [W_{fd}\,p_{1fd}\,p_{3md} + W_{fd}\,p_{3fd}\,p_{1md} + W_{fr}\,p_{3fr}\,p_{1mr}] / \overline{W_f} \\
G''_{22} &= [W_{fd}\,p_{2fd}\,p_{2md} + W_{fr}\,p_{2fr}\,p_{2mr}] / \overline{W_f} \\
G''_{23} &= [W_{fd}\,p_{2fd}\,p_{3md} + W_{fd}\,p_{3fd}\,p_{2md} + W_{fr}\,p_{2fr}\,p_{3mr} + W_{fr}\,p_{3fr}\,p_{2mr}] / \overline{W_f} \\
G''_{33} &= [W_{fd}\,p_{3fd}\,p_{3md} + W_{fr}\,p_{3fr}\,p_{3mr}] / \overline{W_f}
\end{aligned} \qquad (10)
$$

where

$$
\overline{W_f} = \Sigma\, W_{fd}\,p_{i,fd}\,p_{i,md} + \Sigma\, W_{fr}\,p_{i,fr}\,p_{i,mr} \qquad (11)
$$

The mean fitness, $\overline{W}$, is assessed:

$$
\overline{W} = \overline{W_b}\,\overline{W_f} \qquad (12)
$$

Selection is now started on the new generation (3).

This recursion is repeated for $0.5n$ generations in each environment using appropriate c and f values. Although genotype frequencies differ from generation to generation, they remain the same from cycle to cycle, but are not Hardy-Weinberg ratios as erroneously stated in ref. 7.

At the end of the first generation $\overline{W}$ is labelled $\overline{W}_1$, at the end of the second $\overline{W}_2$, etc. At the end of the n generation cycle LGMF is assessed from these mean fitness, $\overline{W}$, values:

$$
LGMF = \sqrt[n]{\overline{W}_1\,\overline{W}_2\,\overline{W}_3 \ldots \overline{W}_n} \qquad (13)
$$

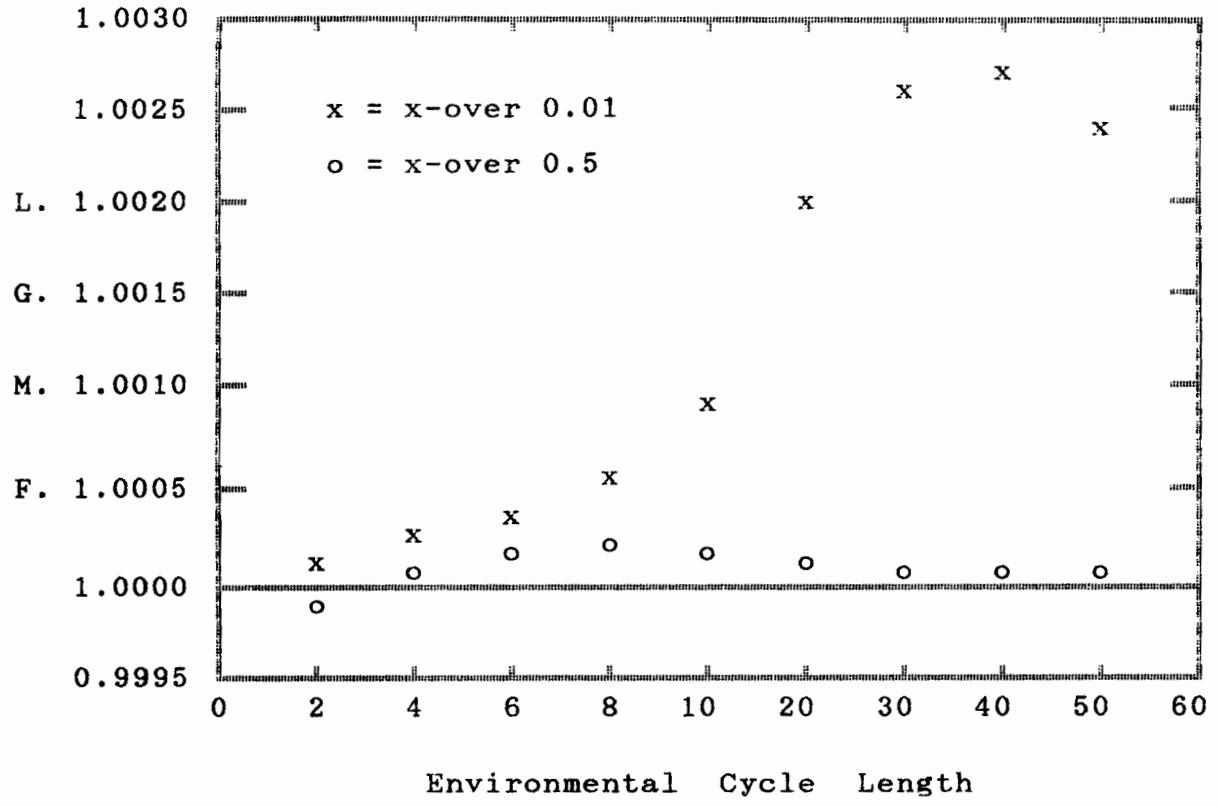

FIG. 1. The *LGMF* of the population when the homosexuality locus is on an autosomal chromosome and the gene is recessive, homosexual, *hh*, males not breeding, $P = 0$. Environmental cycle length is in generations. Mutation is constant, $u = 5 \times 10^{-4}$. In the absence of the gene for homosexuality the *LGMF* = 1.

After several cycles the *LGMF* settles to a repeating value.

The presence of the gene for homosexuality under cycling conditions can raise the *LGMF* of the population. With long environmental period this is achieved by lowering the variance through time in mean fitness. In Fig. 1 the rate of mutation of *H* to *h* is 5×10^{-4}, Lower or higher rates are also favourable. The rate of mutation used in the model must not be so high that the allele for homosexuality becomes relatively common, since the advantage resulting from the chance factor involved in a rare allele being passed on to the next generation would be lost.

Tight linkage is best, but although free recombination gives a small effect, with segregation the locus for homosexuality could interact with all loci affecting viability, and the advantage could be large for cycles longer than two generations.

Under stochastic conditions tight linkage with high mutation rates produces the best *LGMF* results when competing with an all heterosexual population. To study stochastic effects a population with mutation for the homosexuality gene *h* and a population without mutation for *h* are run in parallel, without interbreeding, and their respective *LGMF*s computed.

Homosexuality gene dominant

Homosexual males are of genotype *H-* in this model, *u* being the rate of mutation of *h* to *H*.

Genotypes surviving to breeding time are calculated as in equations (3), $\overline{W}_b$ being calculated as in equation (4).

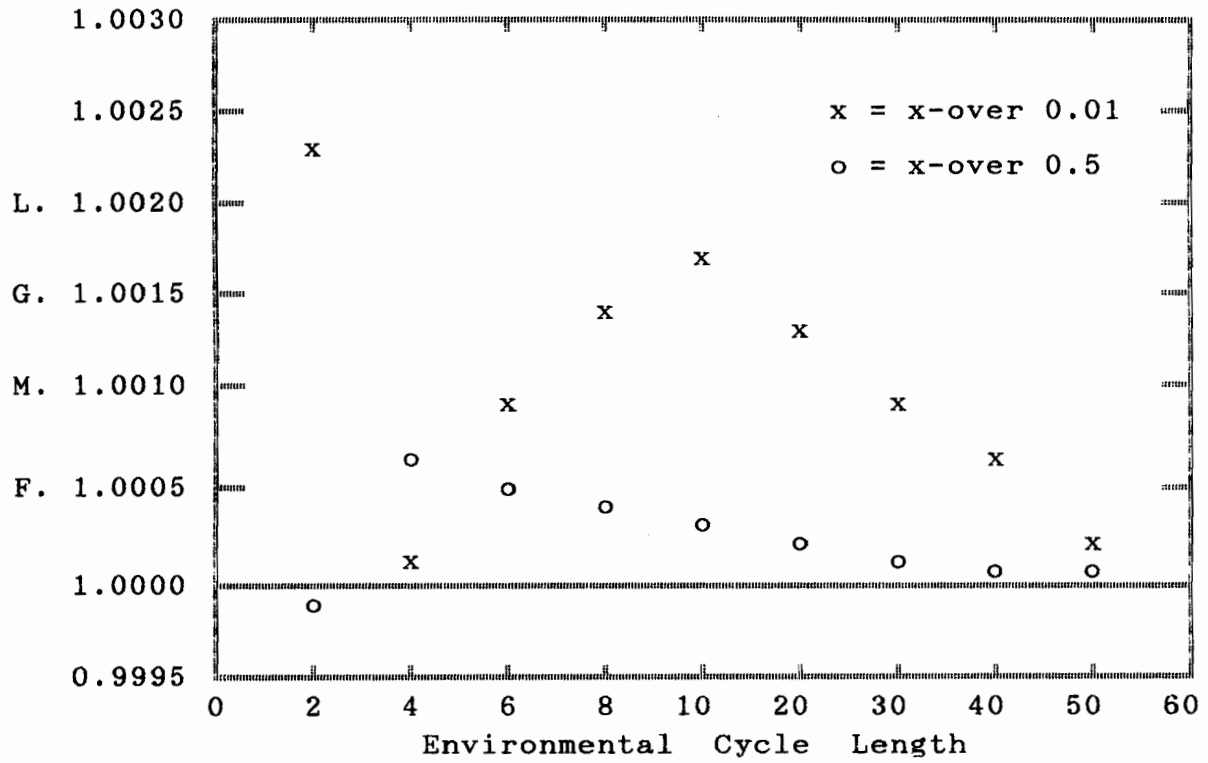

FIG. 2. The *LGMF* of the population when the homosexuality locus is on an autosomal chromosome and the gene is dominant, homosexual, *H-*, males not breeding, $P = 0$. Environmental cycle length is in generations. Mutation is constant, $u = 5 \times 10^{-4}$. In the absence of the gene for homosexuality the *LGMF* = 1.

The mutation episode for the female gives:

$$
\begin{aligned}
G_{00f} &= G'_{00} + uG'_{01} + u^2G'_{11} \\
G_{01f} &= (1-u)G'_{01} + (2u-u^2)G'_{11} \\
G_{11f} &= (1-2u)G'_{11} \\
G_{02f} &= G'_{02} + uG'_{03} + uG'_{12} + u^2G'_{13} \\
G_{03f} &= (1-u)G'_{03} + \tfrac{1}{2}(2u-u^2)G'_{13} \\
G_{12f} &= (1-u)G'_{12} + \tfrac{1}{2}(2u-u^2)G'_{13} \\
G_{13f} &= (1-2u)G'_{13} \\
G_{22f} &= G'_{22} + uG'_{23} + u^2G'_{33} \\
G_{23f} &= (1-u)G'_{23} + (2u-u^2)G'_{33} \\
G_{33f} &= (1-2u)G'_{33}
\end{aligned}
\qquad (14)
$$

Gamete proportions for the female are:

$$
\begin{aligned}
p_{0fd} &= [2G_{00f} + (1-c_d)((1-r)G_{03f}+rG_{12f}+G_{01f}) + G_{02f}] / 2 \\
p_{1fd} &= [2G_{11f} + (1+c_d)((1-r)G_{12f}+rG_{03f}+G_{01f}) + G_{13f}] / 2 \\
p_{2fd} &= [G_{02f} + (1-c_d)((1-r)G_{12f}+rG_{03f})] / 2 \\
p_{3fd} &= [G_{13f} + (1+c_d)((1-r)G_{03f}+rG_{12f})] / 2 \\
p_{2fr} &= [2G_{22f} + (1-c_r)G_{23f}] / 2 \\
p_{3fr} &= [2G_{33f} + (1+c_r)G_{23f}] / 2
\end{aligned} \qquad (15)
$$

Male genotypes breeding are:

$$
S'_m = PG'_{00} + PG'_{01} + PG'_{02} + PG'_{03} + G'_{11} + PG'_{12} + G'_{13} + PG'_{22} + PG'_{23} + G'_{33} \qquad (16)
$$

The mutation episode for the male gives:

$$
\begin{aligned}
G_{00m} &= [PG'_{00} + uPG'_{01} + u^2G'_{11}] / S'_m \\
G_{01m} &= [(1-u)PG'_{01} + (2u-u^2)G'_{11}] / S'_m \\
G_{11m} &= [(1-2u)G'_{11}] / S'_m \\
G_{02m} &= [PG'_{02} + uPG'_{03} + uPG'_{12} + u^2G'_{13}] / S'_m \\
G_{03m} &= [(1-u)PG'_{03} + \tfrac{1}{2}(2u-u^2)G'_{13}] / S'_m \\
G_{12m} &= [(1-u)PG'_{12} + \tfrac{1}{2}(2u-u^2)G'_{13}] / S'_m \\
G_{13m} &= [(1-2u)G'_{13}] / S'_m \\
G_{22m} &= [PG'_{22} + uPG'_{23} + u^2G'_{33}] / S'_m \\
G_{23m} &= [(1-u)PG'_{23} + (2u-u^2)G'_{33}] / S'_m \\
G_{33m} &= [(1-2u)G'_{33}] / S'_m
\end{aligned} \qquad (17)
$$

Gamete proportions for the male are:

$$
\begin{aligned}
p_{0md} &= [2G_{00m} + (1-c_d)((1-r)G_{03m}+rG_{12m}+G_{01m}) + G_{02m}] / 2 \\
p_{1md} &= [2G_{11m} + (1+c_d)((1-r)G_{12m}+rG_{03m}+G_{01m}) + G_{13m}] / 2 \\
p_{2md} &= [2G_{22m} + (1-c_d)((1-r)G_{12m}+rG_{03m}+G_{23m}) + G_{02m}] / 2 \\
p_{3md} &= [2G_{33m} + (1+c_d)((1-r)G_{03m}+rG_{12m}+G_{23m}) + G_{13m}] / 2
\end{aligned} \qquad (18)
$$

p_{0mr}, p_{1mr}, etc. are obtained by replacing c_d with c_r in (18).

Offspring are calculated from equations (10), $\overline{W}_f$ from (11) and $\overline{W}$ from (12).

The *LGMF* is assessed at the end of the cycle, equation (13).

Results (Fig. 2) are very similar to the recessive case, the gene for homosexuality raising the *LGMF* of the population by lowering the variance through in mean fitness.

Homosexuality Gene Locus on the *X* Chromosome

The hypothetical gene for homosexuality at the *H* locus is assumed to be carried on the *X* chromosome only, *h* males being homosexual. First the *A* locus is considered to be carried on an autosomal chromosome, then on the sex chromosomes.

Viability locus on an autosomal chromosome

Female genotypes *XX* surviving to breed are calculated as in equations (3). The mutation episode for the female is as in (5).

Gamete proportions for the female are obtained from equations (6). Segregation is involved here, not recombination, and the value used for r is 0.5.

Male genotypes *XY* surviving to breeding time are assessed, the frequency of the gamete type $A = p_{/0}$ and $a = p_{/2}$, the genotype frequency of $AH/A = G_{0/0}$, $AH/a = G_{0/2}$, $Ah/A = G_{1/0}$ $Ah/a = G_{1/2}$, $aH/A = G_{2/0}$, $aH/a = G_{2/2}$, $ah/A = G_{3/0}$ and $ah/a = G_{3/2}$.

$$\begin{aligned} G'_{0/0} &= W_{bd}\, G_{0/0} / \overline{W}_m & G'_{2/0} &= W_{bd}\, G_{2/0} / \overline{W}_m \\ G'_{0/2} &= W_{bd}\, G_{0/2} / \overline{W}_m & G'_{2/2} &= W_{br}\, G_{2/2} / \overline{W}_m \\ G'_{1/0} &= W_{bd}\, G_{1/0} / \overline{W}_m & G'_{3/0} &= W_{bd}\, G_{3/0} / \overline{W}_m \\ G'_{1/2} &= W_{bd}\, G_{1/2} / \overline{W}_m & G'_{3/2} &= W_{br}\, G_{3/2} / \overline{W}_m \end{aligned} \qquad (19)$$

where

$$\overline{W}_m = \Sigma\, W_b\, G_{i/j} \qquad (20)$$

W_m for the male equals W_b for the female here.

Male genotypes breeding are:

$$S'_m = G'_{0/0} + G'_{0/2} + PG'_{1/0} + PG'_{1/2} + G'_{2/0} + G'_{2/2} + PG'_{3/0} + PG'_{3/2} \qquad (21)$$

The mutation episode for the male gives:

$$\begin{aligned} G_{1/0d} &= [\quad PG'_{1/0} + u(1-c_d)\,G'_{0/0}] / S'_m \\ G_{0/0d} &= [(1-u(1-c_d))\,G'_{0/0}\quad\ ] / S'_m \\ G_{1/2d} &= [\quad PG'_{1/2} + u(1-c_d)\,G'_{0/2}] / S'_m \\ G_{0/2d} &= [(1-u(1-c_d))\,G'_{0/2}\quad\ ] / S'_m \\ G_{3/0d} &= [\quad PG'_{3/0} + u(1-c_d)\,G'_{2/0}] / S'_m \\ G_{2/0d} &= [(1-u(1-c_d))\,G'_{2/0}\quad\ ] / S'_m \\ G_{3/2d} &= [\quad PG'_{3/2} + u(1-c_d)\,G'_{2/2}] / S'_m \\ G_{2/2d} &= [(1-u(1-c_d))\,G'_{2/2}\quad\ ] / S'_m \end{aligned} \qquad (22)$$

u being the rate of mutation of H to h. $G_{0/0r}$ etc, are calculated by replacing c_d with c_r in (22).

Gamete proportions for the male are:

$$\begin{aligned} p_{0md} &= [G_{0/0d} + (1-r)\,G_{0/2d} + rG_{2/0d}] / 2 \\ p_{1md} &= [G_{1/0d} + (1-r)\,G_{1/2d} + rG_{3/0d}] / 2 \\ p_{2md} &= [G_{2/2d} + (1-r)\,G_{2/0d} + rG_{0/2d}] / 2 \\ p_{3md} &= [G_{3/2d} + (1-r)\,G_{3/0d} + rG_{1/2d}] / 2 \end{aligned} \qquad (23a)$$

$$\begin{aligned} p_{/0m} &= [G'_{0/0} + (1-r)\,G'_{2/0} + rG'_{0/2} + PG'_{1/0} + (1-r)\,PG'_{3/0} + rPG'_{1/2}] / 2S'_m \\ p_{/2m} &= [G'_{2/2} + (1-r)\,G'_{0/2} + rG'_{2/0} + PG'_{3/2} + (1-r)\,PG'_{1/2} + rPG'_{3/0}] / 2S'_m \end{aligned} \qquad (23b)$$

p_{0mr} etc. are calculated by using $G_{0/0r}$ etc. in (23a). Equation (23a) represents X gametes and (23b) Y gametes, r = 0.5.

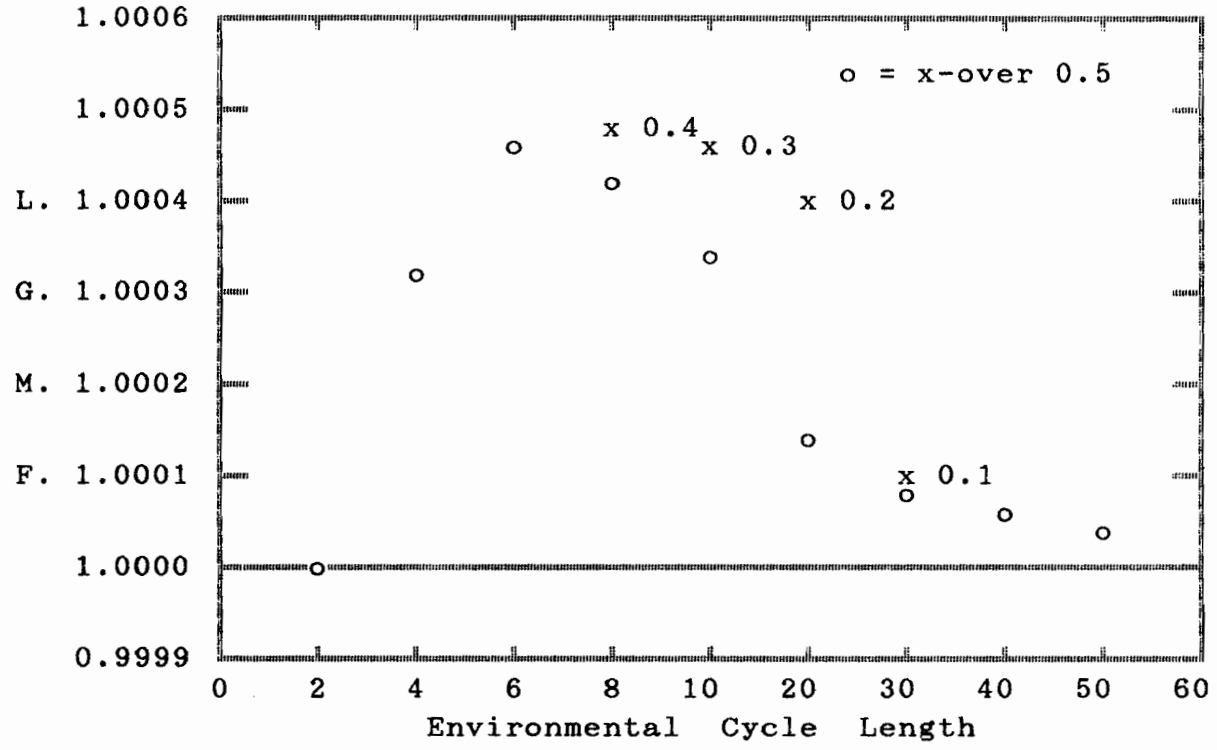

FIG. 3. The *LGMF* of the population when the homosexuality locus is on the X chromosome, homosexual, h, males do not breed, P = 0. Segregation, not recombination, is involved when the viability locus is carried on an autosomal chromosome. Linkage is involved when the viability locus is carried on the sex, i.e. X and Y, chromosomes, optimum linkage values being given on the figure (the gene for homosexuality is only advantageous if environmental period >6 and <40). The mutation rate is constant, $u = 5 \times 10^{-4}$. In the absence of the gene for homosexuality the *LGMF* = 1.

Female offspring are calculated from equations (10).
Male offspring are:

$$
\begin{aligned}
G''_{0/0} &= [W_{fd}\, p_{0fd}\, p_{/0m} \qquad\qquad\quad] \,/\, \overline{W_f} \\
G''_{0/2} &= [W_{fd}\, p_{0fd}\, p_{/2m} \qquad\qquad\quad] \,/\, \overline{W_f} \\
G''_{1/0} &= [W_{fd}\, p_{1fd}\, p_{/0m} \qquad\qquad\quad] \,/\, \overline{W_f} \\
G''_{1/2} &= [W_{fd}\, p_{1fd}\, p_{/2m} \qquad\qquad\quad] \,/\, \overline{W_f} \\
G''_{2/0} &= [W_{fd}\, p_{2fd}\, p_{/0m} + W_{fr}\, p_{2fr}\, p_{/0m}] \,/\, \overline{W_f} \\
G''_{2/2} &= [W_{fd}\, p_{2fd}\, p_{/2m} + W_{fr}\, p_{2fr}\, p_{/2m}] \,/\, \overline{W_f} \\
G''_{3/0} &= [W_{fd}\, p_{3fd}\, p_{/0m} + W_{fr}\, p_{3fr}\, p_{/0m}] \,/\, \overline{W_f} \\
G''_{3/2} &= [W_{fd}\, p_{3fd}\, p_{/2m} + W_{fr}\, p_{3fr}\, p_{/2m}] \,/\, \overline{W_f}
\end{aligned}
\qquad (24)
$$

W_f is calculated from female (11) or male offspring totals, which are identical. The *LGMF* is assessed at the cycle end (13).

The presence of the rare homosexuality gene *h* raises the *LGMF* of a population if cycles are longer than two generations (Fig. 3). Under stochastic conditions higher mean *LGMF* values are obtained than for an all heterosexual population. It is to be noted that the gene for homosexuality can interact with all loci affecting viability, segregation being involved.

Viability locus on the sex, *X* and *Y*, chromosomes

This highly unlikely situation is investigated out of interest. The same model as above is used, except linkage is involved. The hypothetical gene for homosexuality on the *X* chromosome is assumed to be in the *Xq*28 region near the end of the long arm[3]. The viability locus is carried on both the *X* chromosome, in the *DXY*5154 region at the end of the long arm, and on the *Y* chromosome. Tight linkage is to be expected since recombination in this region is low. The optimal recombination rates are given in Fig. 3., as is apparent they are not very low, tight linkage is unfavourable, the *LGMF* being lowered.

Viability locus on the *X* chromosome only

Female genotypes *XX* surviving to breed are calculated from equations (3). The mutation episode for the female is as in (5).

Gamete proportions for the female are obtained from (6).

Male genotypes *XY* surviving to breeding time are calculated, the frequency of the genotype *AH/Y* = $G_{0/Y}$, *Ah/Y* = $G_{1/Y}$, *aH/Y* = $G_{2/Y}$ and *ah/Y* = $G_{3/Y}$, *Y* being the *Y* chromosome:

$$\begin{aligned} G'_{0/Y} &= W_{bd}\, G_{0/Y} / \overline{W_m} \\ G'_{1/Y} &= W_{bd}\, G_{1/Y} / \overline{W_m} \\ G'_{2/Y} &= W_{br}\, G_{2/Y} / \overline{W_m} \\ G'_{3/Y} &= W_{br}\, G_{3/Y} / \overline{W_m} \end{aligned} \qquad (26)$$

where

$$\overline{W_m} = \Sigma\, W_b\, G_{i/Y} \qquad (27)$$

W_m for the male equals W_b for the female when b = 1, i.e. $W_b = 1$.

Male genotypes breeding are:

$$S'_m = G'_{0/Y} + PG'_{1/Y} + G'_{2/Y} + PG'_{3/Y} \qquad (28)$$

The mutation episode for the male results in the following gamete proportions, no recombination or segregation is involved:

$$
\begin{aligned}
p_{0md} &= [(1-u(1-c_d))\,G'_{0/Y} \qquad\qquad\quad] \,/\, 2S'_m \\
p_{1md} &= [\qquad\quad PG'_{1/Y} + u(1-c_d)\,G'_{0/Y}] \,/\, 2S'_m \\
p_{2md} &= [(1-u(1-c_d))\,G'_{2/Y} \qquad\qquad\quad] \,/\, 2S'_m \\
p_{3md} &= [\qquad\quad PG'_{3/Y} + u(1-c_d)\,G'_{2/Y}] \,/\, 2S'_m
\end{aligned} \tag{29a}
$$

$$
p_Y = [G'_{0/Y} + PG'_{1/Y} + G'_{2/Y} + PG'_{3/Y}] \,/\, 2S'_m \tag{29b}
$$

p_{0mr} etc. are calculated by replacing c_d with c_r in (29a). Equation (29a) represents X gametes and (29b) Y gametes.

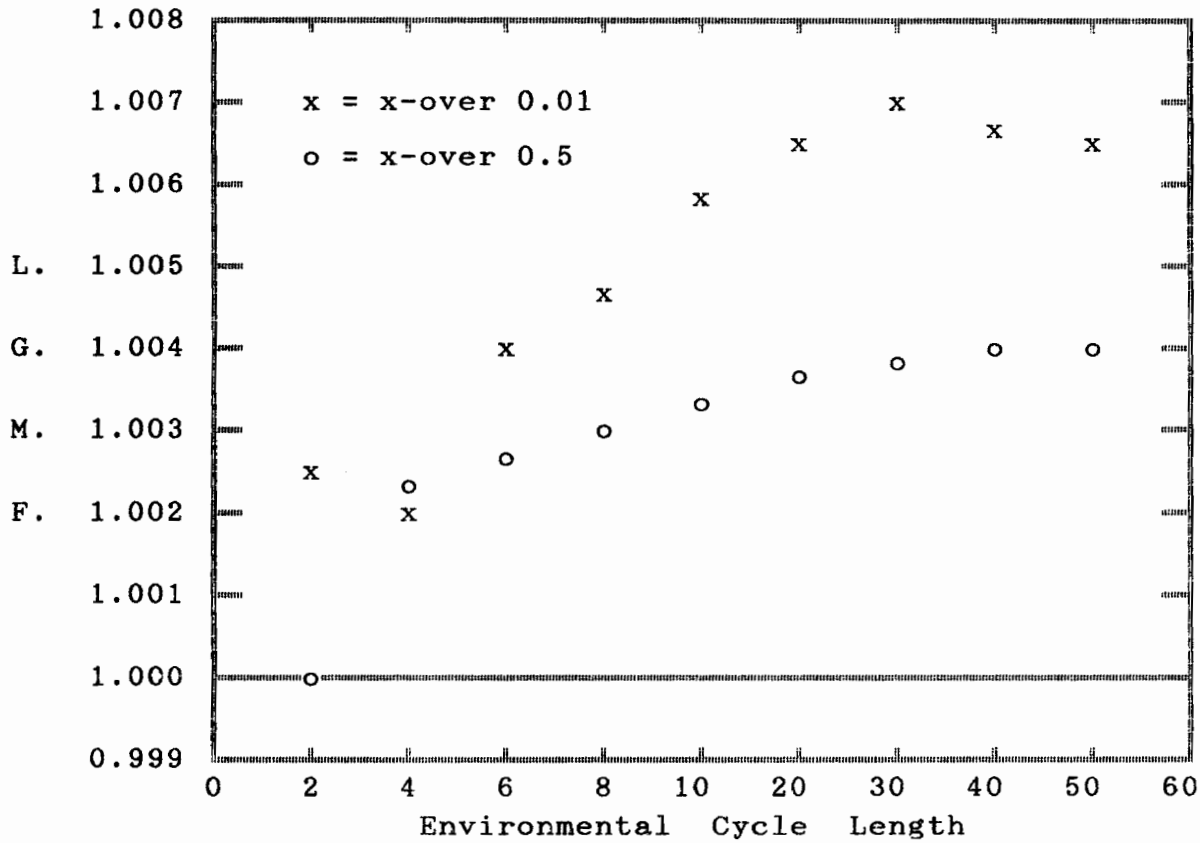

FIG. 4. The *LGMF* of the population when the viability locus and the homosexuality locus are both on the X chromosome, homosexual, h, males do not breed, $P = 0$. The rate of mutation is constant, $u = 5 \times 10^{-4}$. In the absence of homosexuality the *LGMF* = 1.

Female offspring are calculated from equations (10).
Male offspring are:

$$
\begin{aligned}
G''_{0/0} &= [W_{fd}\,p_{0fd}\,p_Y \qquad\quad] \,/\, \overline{W_f} \\
G''_{1/0} &= [W_{fd}\,p_{1fd}\,p_Y \qquad\quad] \,/\, \overline{W_f} \\
G''_{2/0} &= [W_{fd}\,p_{2fd}\,p_Y + W_{fr}\,p_{2fr}\,p_Y] \,/\, \overline{W_f} \\
G''_{3/0} &= [W_{fd}\,p_{3fd}\,p_Y + W_{fr}\,p_{3fr}\,p_Y] \,/\, \overline{W_f}
\end{aligned} \tag{30}
$$

$\bar{W}_f$ is calculated from female (11) or male offspring totals, which are identical. The *LGMF* is computed at the cycle end (13).

The advantage is relatively strong (Fig. 4) where the homosexuality locus is linked to the viability locus on the *X* chromosome. Linkage always raises the *LGMF* above 1, as does free recombination if cycle length is longer than two generations. Under stochastic conditions the population with the homosexuality gene *h* performs better than an all heterosexual population.

Conclusion

The results indicate that the hypothetical gene for male homosexuality can raise the long term geometric mean fitness, *LGMF*, of a population by lowering its variance through time in mean fitness. This buffering effect is particularly advantageous when environmental cycles are long, the population showing smaller fluctuations in size. The gene for homosexuality at the *H* locus in the models presented tends to become associated with the currently fitter allele at any general viability, *A*, locus. Fitter families tend to be larger and are more likely to have the currently fittest viability allele, but they also have a greater chance of passing on the mutant homosexuality gene to their offspring. This damps down oscillations in gene frequencies, the *A* locus allele being lost along with the homosexuality gene at the *H* locus if the male homosexual does not breed.

The greatest effect of the hypothetical gene for homosexuality is given when the environmental period is around 20 to 50 generations and the locus is carried on the *X* chromosome along with a general viability locus, or it is linked to a general viability locus on an autosomal chromosome and the gene for homosexuality is recessive. This environmental period corresponds in the case of overlapping generations to the generation time of man of 30 years. The advantage as regards long term geometric mean fitness to the population is in the order of 2,000 to 7,000 individuals in one million. If several loci for homosexuality linked to viability loci on autosomal chromosomes were involved, the advantage would be even greater. With a single locus for homosexuality, and free recombination or segregation, the advantage is in the order of 200 individuals in one million, not a negligible figure if all loci affecting viability were included when it could mount into the thousands.

With this long environmental period, or generation time in the case of man, the gain in longterm geometric mean fitness is due to the more stable population size achieved by the reduction in variance through time in mean fitness when some homosexual males do not breed. In other words the population fluctuates less in size than an all heterosexual population.

If homosexuality were genetically determined along the lines of the models presented, the advantage in terms of population dynamics in the context of the developed world would not be that male homosexuality increased population size, but that it helped stabilise the size of the population. In the absence of the buffering effect of homosexuality, it is conceivable that lethal and sub-lethal genes might increase in frequency. These genes can play a similar role in population dynamics but, apart from being detrimental, they are not so effective[7].

The possibility that homosexuality genes may possess superior fitness in heterozygous conditions[6] would not apply to the present findings. The advantage found depends on the probability of a rare gene for homosexuality being lost each generation. For this chance factor to act, continuous mutation is necessary.

In the models the frequency of the mutant gene diminishes by about one half each generation, and is rapidly lost, much faster than predicted by simple Mendelian genetics. The same phenomenon is observed with rare lethal and sub-lethal genes, these being lost from populations at similar rates, much faster than predicted for a lethal recessive gene by simple Mendelian genetics (see Sturtevant in ref. 15). This is probably due to the chance factor involved in heterozygous carriers passing on the gene[7]. Whether the mutant gene for homosexuality is recessive or dominant or carried on the X chromosome makes little difference to the rate of loss, neither does the percentage of male homosexuals breeding when this is less than 90%.

The idea that a gene for male homosexuality could be kept in circulation by being carried through the female line as the result of a heterozygote advantage to carriers is unrealistic. If overdominance were responsible[6], a female heterozygous carrier of the gene would have to be very much fitter than other females. Of course a survey might give this impression if larger families had homosexual sons more frequently than expected, since the chance factor involved in the gene being passed on to progeny is reliant on family size anyway. If a contribution of the homosexual male to the benefit of his siblings, or the families of his siblings, were responsible this likewise would have to be very large to maintain the gene in the population.

The frequency of homosexual males in human populations is considered to be around 2% by conservative estimations[3]. The mutation rate had to be 5×10^{-4} to maintain homosexual males at 2% in the population when the hypothetical gene for homosexuality was carried on the X chromosome. When the gene for homosexuality was on an autosomal chromosome, higher mutation rates were necessary to maintain homosexuals at 2%, a mutation rate of 5×10^{-4} only maintaining homosexuals at a frequency around 0.5%. This could signify that several loci for homosexuality are

involved, as is indicated by the results. That several genes, or other factors, are also involved in homosexuality is expected[3].

The finding that recurrent mutation would be necessary to maintain homosexual males in a population if homosexuality were genetically determined indicates interference with the genetic code, apart from being unethical, would be ineffectual in lowering the frequency of homosexuality.

Human sexual orientation appears to be genetically diverse and, to quote Mary-Claire King[16], "to identify variation and characterize genetic influences on it is to recognize the significance and value of that variation, not to condemn it".

ACKNOWLEDGEMENTS

I thank B.J. Cooper for comments on the original manuscript, and S. Orzak and an anonymous referee for helpful criticism of an earlier version of the paper.

LITERATURE CITED

1. Pillard, R.C. & Weinrich, J.D. 1986. *Arch. Gen. Psychiatry* **43** : 808-812.
2. Bailey, J.M. & Pillard, R.C. 1991. *Arch. Gen. Psychiatry* **48** : 1089-1096.
3. Hamer, D.H., Hu, S., Magnuson, V.L., Hu, N. & Pattatucci, A.M.L. 1993. *Science* **261** : 321-327.
4. Hu, S., Pattatucci, A.M.L., Patterson, C., Li, L., Fulker, D.W., Cherry, S.S., Kruglyak, L. & Hamer, D.H. 1995. *Nature Genetics* **11** : 248-256.
5. Dawkins, R. 1982. *The Extended Phenotype* (W.H. Freeman & Co. Ltd).
6. Hutchinson, G.E. 1959. *Am. Nat.* **93** : 81-91.
7. Cook, M. 1995. *Evol. Theory* **10** : 283-297.
8. Endler, J.A. 1986. *Ox. Surv. Evol. Biol.* **3** : 224-243.
9. Strickberger, M.W. 1985. *Genetics* (Collier Macmillan Canada Incorporated).
10. Fisher, R.A. 1930. *The Genetical Theory of Natural Selection* (Clarendon, Oxford; reprinted 1958, Dover, New York).
11. Gillespie, J.H. 1974. *Genetics* **76** : 601-606.
12. Stearns, S.C. 1986. in *Patterns and Processes in the History of Life.* (eds. D.M. Raup & D. Jablonski. Springer Verlag, Berlin.)
13. Cook, M. 1992. *Evol. Theory* **10** : 129-143.
14. Cook, M. 1995. *Evol. Theory* **11** : 33-50.
15. Dobzhansky, Th. & Wright, S. 1941. *Genetics* **26** : 23-51.
16. King, M-C. 1993. *Nature* **364** : 288-289.